Claims

What is claimed is:

1. A method for scheduling responses in a point-to-point communication system having *k* channels, the method comprising the steps of:

receiving at a central server a plurality of job requests; and

determining an adaptive schedule for transmitting, via said *k* channels, data responsive to said job requests, wherein the servicing of a first job request via a first channel is interrupted, an unserviced portion of said data is returned from a local channel server to said central server, and said unserviced portion is subsequently serviced via a second channel so as to service a second job request via said first channel in accordance with an updated schedule.

2. The method of claim 1, further comprising the step of servicing said job requests via a plurality of local channel servers in accordance with said adaptive schedule.

3. The method of claim 1, wherein said step of determining an adaptive schedule further comprises the step of updating said schedule upon the arrival of a new job request at said central server.

4. The method of claim 1, wherein said step of determining an adaptive schedule further comprises the step of updating said schedule upon the arrival at said central server of said interrupted first job request.

5. The method of claim 1, further comprising the step of calculating a deadline, D_i, for each of said plurality of job requests, given by $D_i = S \times P_i + A_i$, wherein P_i is a processing time corresponding to each said job request, A_i is an arrival time corresponding to each said job request, and S is a stretch value, wherein said schedule for processing said job requests is adaptively

determined such that each said job request is serviced via said channels within said deadline corresponding to each said job request.

6. The method of claim 5, wherein said processing time further comprises a value corresponding to the size of each said job request divided by a channel bandwidth.

7. The method of claim 5, further comprising the step of proposing said stretch value.

8. The method of claim 7, further comprising the step of calculating said deadlines for each of said job requests as a function of said proposed stretch value.

9. The method of claim 8, wherein said scheduling step further comprises scheduling said job requests, at said proposed stretch value, according to an "earliest deadline first" arrangement.

10. The method of claim 9, wherein said stretch value is a feasible stretch value if, when processed in an order dictated by said "earliest deadline first" arrangement, a completion time of each and every said job request is prior to said corresponding deadline for each said job request at said feasible stretch value.

11. The method of claim 10, further comprising the steps of:

iteratively adjusting said stretch value until said stretch value is a feasible stretch value; and

after finding said feasible stretch value, processing said job requests in an order dictated by said "earliest deadline first" arrangement, said "earliest deadline first" arrangement corresponding to said stretch values determined in the latest iteration.

12. The method of claim 11, wherein said adjusting step further comprises adjusting said proposed stretch value by doubling said proposed stretch value.

13. The method of claim 12, wherein said feasible stretch value further comprises an optimal feasible stretch value, said method further comprising the step of adjusting said proposed stretch value until a feasible stretch value is found, said optimal feasible stretch value corresponding to the smallest feasible stretch value that permits said each and every job request to have a completion time prior to said deadline of said job request at said optimal feasible stretch value.

14. The method of claim 13, wherein said adjusting step further comprises performing a binary search for said optimal feasible stretch value after said controller determines that said feasible stretch value falls between said proposed stretch value and double said proposed stretch value.

15. The method of claim 5, wherein said plurality of job requests further comprises a set having n number of job requests, wherein n is an integer.

16. The method of claim 15, wherein said integer n has a value that is equal to a total number of job requests processed by said server system.

17. The method of claim 15, wherein said integer n has a value that is less than a total number of job requests processed by said server system.

18. The method of claim 1, further comprising the step of processing a plurality of pending job requests prior to receiving at said central server said new job request, wherein said communication system is an on-line system.

19. The method of claim 18, further comprising the step of calculating a deadline, D_i, for each of said received and pending job requests, given by $D_i = S^* \times P_i + A_i$, wherein P_i is a processing time corresponding to each said job request, A_i is an arrival time corresponding to each said job request, and S^* is a modified stretch value estimated based on job requests arrived during a history window defined by a history window parameter.

20. The method of claim 19, wherein said processing time further comprises a value corresponding to the size of each said job request divided by a channel bandwidth.

21. The method of claim 19, wherein said modified stretch value corresponds to a maximum stretch value for all of said job requests arrived within said history window.

22. The method of claim 19, wherein said scheduling step further comprises scheduling said job requests according to an "earliest deadline first" arrangement, said "earliest deadline first" arrangement utilizing said deadlines calculated as a function of said modified stretch value.

23. The method of claim 22, further comprising the step of adjusting said stretch value upon an arrival of a new job request.

24. The method of claim 22, further comprising the step of adjusting said stretch value upon said completion of any said job request.

25. An apparatus for use in scheduling responses in a point-to-point communication system having k channels, the apparatus comprising:

a central server configured: (i) to receive a plurality of job requests; and (ii) to determine an adaptive schedule for transmitting, via said k channels, data responsive to said job requests, wherein the servicing of a first job request via a first channel is interrupted, an unserviced portion of said data is returned from a local channel server to said central server, and said unserviced portion is subsequently serviced via a second channel so as to service a second job request via said first channel in accordance with an updated schedule.

26. An article of manufacture for storing one or more software programs for use in a central server for scheduling responses in a point-to-point communication system having k channels, wherein the one or more programs when executed implement the step of:

determining an adaptive schedule for transmitting, via said k channels, data responsive to received job requests, wherein the servicing of a first job request via a first channel is interrupted, an unserviced portion of said data is returned from a local channel server to said central server, and said unserviced portion is subsequently serviced via a second channel so as to service a second job request via said first channel in accordance with an updated schedule.